# Viewing Student Affect and Learning through Classroom Observation and Physical Sensors

Toby Dragon,[1] Ivon Arroyo,[1] Beverly P. Woolf,[1] Winslow Burleson,[2]
Rana el Kaliouby,[3] Hoda Eydgahi[4]

[1] Department of Computer Science, University of Massachusetts-Amherst
[2] Arts, Media and Engineering Program, Arizona State University
[3] Media Lab, Massachusetts Institute of Technology
[4] Department of Electrical Engineering, Massachusetts Institute of Technology
{dragon, arroyo, bev}@cs.umass.edu; winslow.burleson@asu.edu;
kaliouby@media.mit.edu; hoda@mit.edu

**Abstract.** We describe technology to dynamically collect information about students' emotional state, including human observation and real-time multi-modal sensors. Our goal is to identify physical behaviors that are linked to emotional states, and then identify how these emotional states are linked to student learning. This involves quantitative field observations in the classroom in which researchers record the behavior of students who are using intelligent tutors. We study the specific elements of learner's behavior and expression that could be observed by sensors. The long-term goal is to dynamically predict student performance, detect a need for intervention, and determine which interventions are most successful for individual students and the learning context (problem and emotional state).

## 1 Introduction and Previous Work

The obvious next frontier in computational instruction is to systematically examine the relationship(s) between student affective and learning outcome (performance) [18]. Human emotion is completely intertwined with cognition in guiding rational behavior, including memory and decision-making [18,11,16,5]. Students' emotion towards learning can have a drastic effect on their learning experience [10]. An instructor who establishes emotional and social connections with a student in addition to cognitive understanding enhances the learning experience. Responding to a learner's emotion, understanding her at a deep level, and recognizing her affect (e.g. bored, frustrated or disengaged) are key elements of quality teaching. If computer tutors are to interact naturally with humans, they need to recognize affect and express social competencies. This research attempts to understand how students express emotion, detect these emotions, and quantify emotional variables.

Previous projects have produced computational tutors that recognized and responded to models of emotion (e.g., self-efficacy and empathy [15]). Projects have tackled the sensing and modeling of emotion in learning and educational gaming environments [14, 17]. A dynamic decision network was used to measure student emotional state based on variables such as heart rate, skin conductance and eyebrow position [7]. Studies have evaluated the impact of affective interface agents on both affective and motivational outcomes based factors (e.g., gender, ethnicity). Lack of engagement was shown empirically to correlate with a decrease in learning [4]. In this study, however the tutor elicited negative feelings from students, in part because it blocked those who were presumed to be gaming the system [1]. Most prior work on emotion recognition has focused on deliberately expressed emotions within a laboratory setting and not in natural situations such as classroom learning. Many of earlier systems did not use fully adaptive learning environments and some were games. The research described here takes the next step by integrating emotion detection within an intelligent tutor as part of learning in a natural classroom setting.

## 2 Overall Plan

The long-term goal of this research is to dynamically collect information about students' emotional state in order to predict performance, detect a need for intervention, and determine which interventions are most successful for individual students and context (problem, emotional state). To accomplish these tasks, we implement emotion detection within an existing tutor in three phases: classroom observations, the use of physiologic sensors, and software algorithms (e.g., machine learning). We triangulate among these approaches to resolve toward agreement (with the realization that we may be far away from realizing any consensual agreement). This paper describes the first two methods for detection of emotion; classroom observations and a sensor platform.

In the first phase of this research human observation in the classroom approximated the type of information the sensors would collect, and corroborated what sensor information indicates about students' emotional state. Classroom observations are a useful exploratory strategy because human observers can intuitively discern high-level behaviors and make appropriate judgments on limited information that may be difficult to automatically decide from raw sensor data.

In the second phase we evaluate low-cost portable and readily deployable sensors that dynamically detect emotion using the theoretical basis formed from classroom observations. Sensors are can collect constant streams of data in parallel, allowing for much more consistent observation than a human ever could accomplish. They are also increasingly inexpensive and fast at processing/collecting data. Thus, human observations identify behaviors that are worth observing and then sensors gather this behavioral data in bulk. We will evaluate the effectiveness of sensors in predicting student emotional state, and use reinforcement-learning techniques to decide which interventions are most successful for students in certain emotional states.

## 3 Classroom Observations

Our goal in the first phase of this research was to observe student behavior and identify variables that represented 1) emotions and desirable/undesirable states linked to student learning, and 2) physical behaviors linked to emotion states. This involved quantitative field observations in the classroom in which researchers recorded the behavior of students using intelligent tutors. Observations by multiple observers, using this method, have had high inter-rater reliability and report relatively low impact on student behavior once students are used to the observer's presence [4]. Researchers observed students using the Wayang Mathematics Tutor, a tutor that prepares 12-16 year old students for the mathematics section of standardized exams [2]. The tutor, which has been used by a thousand of students represents mathematic skills and recognizes which skills a student has learned. It shows students their progress and offers them a choice of problem difficulty.

### 3.1 Experimental Design

The study included thirty four (34) students in a public school in urban Holyoke, MA, split into 3 different classes. Students took a pretest survey to evaluate their attitudes towards math (self-concept and value) and goal (learning vs. performance) orientation [10], as well as a mathematics pretest with multiple problems to evaluate diverse concepts taught within the Wayang Outpost math tutoring software. Students used the tutoring software during a period of 3 weeks and were then given a posttest. While students used the Wayang software, three researchers coded behavioral variables and subjective variables, such as valence of the student's emotion. Researchers were trained during several sessions to code these variables by observing videos of students using Wayang. Coders rotated around the classroom, coding one student at a time. Observation periods lasted for approximately 15 seconds, with the following 15 seconds to confirm the observation. Because students may have experienced several behaviors/emotions during one time period (e.g., the student was seen forward and then back on the chair), we coded the first state seen, but the second one was coded and taken account later in the analysis.

**Behavioral and Task-Based Variables.** Researchers coded physical behavior (chair and head posture, movement, face gestures) and looked for expressed affect in specific facial expressions (smile, frown, nod) and verbal behavior (loud comments, talk with others). They also coded whether a student appeared to be on- or off-task. The process of identifying this behavior is obviously somewhat subjective and noisy (i.e. a student may look to be on task when they are not). Students were marked as being off-task when they were clearly not using the software appropriately. This includes not looking at the screen, using other programs on the computer, staring blankly at the screen without taking any action, conversing with peers about other subject matter, etc [4]. On-task students might be reading/thinking about the problem, talking to a friend about the problem, or writing a solution on paper. Off-task students are not concentrated/engaged on learning and this is undesirable for learning.

**Emotional Indicators.** Because it is often difficult to distinguish one emotion from another, we limited the conventional emotional terms to four categories of emotions that result from the combination of two indicators: (i) *valence* (positive or negative nature of the emotion/energy the student seemed to be expressing) and (ii) *arousal* or level of physical activity. These emotion indicators are used to express the four basic emotions in Table 1, and are consistent with early research on emotions [20]. However, our concern was that this emotional state variable might not be correlated to learning without also considering on-task or off-task behavior. It is highly desirable for a student to experience a state of joy/excitement when she is on-task, but if the student tends to be joyful while off-task, the emotion variable will not correlate strongly with optimal learning. Thus, we created another variable, *Desirability Value*, which is both task- and emotion-dependent (on/off-task, valence and arousal), see Table 1. The values reflect the fact that being off-task is undesirable, but also that being tired/bored (negative valence, negative arousal) while being on-task is also not desirable, as the student may give up. Frustration while being on-task is not necessarily negative; learning episodes often have productive moments of frustration. Finally, states of positive valence while being on-task are highly desirable, whether accompanied by high arousal or by low levels of arousal where students experience high mental activity without significant observable emotional expression.

**Table 1.** Desirable State Variables and Possible Emotion Indicators

| Valence | Arousal | On/Off task | Example Student Behavior | Desirability value | |
|---|---|---|---|---|---|
| + | + | On | Aha moment, yes! That's it! | 2 | Highly Desirable |
| + | — | On | Concentrated on problem-solving | 2 | Highly Desirable |
| — | + | On | Frustrated with tutoring software, | 1 | Maybe desirable |
| — | — | On | Yawning, zoned out within software | 0 | Not desirable |
| + | + | Off | Laughing with friend | 0 | Not desirable |
| + | — | Off | Very focused but on other software | 0 | Not desirable |
| — | + | Off | Angry quarrel with friend | 0 | Not desirable |
| — | — | Off | Zoned out, or sleeping | 0 | Not desirable |

### 3.2 Results

We evaluated correlations among the frequency of behaviors, task and emotional state variables. Correlations were computed between global emotion indicators and intermediate emotion/task-based state variables. Then we analyzed the correlation between these state-based variables and student behaviors. Students were detected to be on-task 76% of the time, slightly lower than previous findings regarding off/on-task behavior with software learning environments [3].

**Table 2.** Frequency of Emotion Indicators and Desirable Learning States

| Emotion indicators: Valence & Arousal | Frequency | Percent |
|---|---|---|
| + valence & --arousal (concentrated, satisfied) | 148 | 58% |
| + valence & + arousal (excited, joyful, actively engaged) | 85 | 34% |
| - valence & +arousal (frustrated, angry) | 16 | 6% |
| - valence & --arousal (bored, tired) | 5 | 2% |
| Total | 254 | 100% |

| Desirable State | Frequency | Percent |
|---|---|---|
| Highly desirable | 181 | 73% |
| Not desirable | 61 | 24% |
| Medium Desirable | 7 | 3% |

Table 2 shows the frequencies of different emotional states. Note that negative valence emotions were observed only 8% of the time. This could be largely due to the fact that a neutral or indiscernible valence was coded as positive. Table 2 shows that 73% highly desirable states were observed, 3% medium desirable states, and 24% non-desirable states.

**Correlations Between Emotion Indicators and Learning/Attitudes.** We analyzed whether we can use emotional indicators and other state variables to predict learning and motivation, the variables we want to optimize

*Valence*. Valence (or student energy) was significantly correlated to pretest math score (N=34, R=.499, p=.003). This suggests that students who are good in math to begin with, also have substantially more positive emotions while using the software, or at least less unpleasant emotions (e.g. boredom, frustration). Valence was also positively correlated to posttest learning orientation (N=30, R=.499, p<.01), but not to pretest learning orientation, suggesting that having positive valence during the tutoring session may instill higher *learning orientation* goals at posttest time. A similar effect happened for posttest self-concept and valence (R=.48, p<0.01) where students who had higher valence emotions had higher posttest self-concept scores. Thus, the presence of *positive* or *negative* emotions can help predict more general attitudes towards math at posttest time.

*Arousal*. Arousal (or student activity) was negatively correlated with pre-tutor learning orientation (N=30, R=-.373, p<0.05), suggesting that students who are *performance-oriented* (characterized by a desire to be positively evaluated by others) are more likely to be physically active or 'aroused', as opposed to those who are *learning oriented,* who tend to express less physical activity.

*Emotion (Valence + Arousal).* Our emotional scale was correlated with pretest self-concept (R=.385, p<0.05) and posttest learning orientation (R=.463, p<.05), suggesting that the presence of four types of emotions (determined by combinations of valence and arousal) can help predict more general attitudes towards learning math.

*On/Off task.* Being on-task is significantly correlated to posttest self-concept in math (N=30, R=.442, p=.02), but not to pretest self-concept in math, suggesting that being on-task is not a result of an incoming high self-concept in math. However, it indicates that being on-task may generate better self-concept after using the tutor. There is a significant correlation between math posttest performance and being on-task (R=.640, p<.018). Again, being on-task is not correlated with math pretest performance, meaning that prior math knowledge will not predict students' tendencies towards on or off-task behavior. Instead, being on-task seems to lead to higher posttest scores, again implying that being engaged with the tutoring system is part of the reason for achieving higher posttest scores. This is consistent with past research results on on/off task behavior [3]. If we can encourage students to be on-task, we will foster better attitudes for math and higher posttest scores.

*Desirable Learning State***.** Similar significant correlations were found for this variable as on/off task (i.e., it predicted posttest scores and posttest self-concept in math to a similar extent as on/off task behavior). If we can encourage students to be in our desirable learning states (Table 1), we will also foster better attitudes for math and higher posttest scores.

**Correlations Between Emotional/Task-Based States and Behavior.** Several correlations were discovered among student behavior (chair, head and hand position), emotion indicators (valence and arousal) and the desirability value, see Table 3. Clearly, a high positive correlation exists for arousal and chair movement since we defined arousal by physical activity. Meanwhile, valence is not linked to chair movement, meaning that students do not express their positive or negative emotions with chair movement. A negative correlation exists for desirable state and being on-task, meaning that students are in a more desirable learning state (and more on-task) when they don't move so much in the chair.

Other interesting findings (some not shown) are that students with positive valence emotions tend to sit in the middle of the chair, instead of being towards the side, the front or the back of the chair. Last, students leaning on their hands correlated negatively with arousal –as leaning is a fairly inactive posture. It is not that obvious though that students in a state of positive valence also tend to lean on their hands.

**Table 3.** Pearson correlations among student behavior (chair, head and hand position), emotion indicators (valence and arousal), the desirability value and student talk.

| | VALENCE | AROUSAL | ON TASK? | Desirability Value | TALK |
|---|---|---|---|---|---|
| Chair Movement<br>N = | -.467 (0.46*)<br>252 | .420 (.000***)<br>252 | -.140 (.027*)<br>249 | -.154 (.015*)<br>247 | ----- |
| CHAIR Middle<br>N = | .148 (.018*)<br>252 | .107 (090)<br>252 | -.002 (.974)<br>249 | -.003 (.967)<br>247 | ----- |
| HEAD MOVE | -.224 (.000***)<br>249 | .345 (.000***)<br>249 | -.417 (.000***)<br>246 | -.435 (.000***)<br>244 | |
| HEAD SIDE | -.195 (.002**)<br>254 | .247 (.000***)<br>254 | -.325 (.000***)<br>251 | -.337 (.000***)<br>249 | ----- |
| HEAD MOVE SIDE<br>N = | -.270 (.000***)<br>249 | .230 (.000***)<br>249 | -.422 (.000***)<br>246 | -.443 (.000***)<br>244 | ----- |
| HEAD MIDDLE<br>N = | .202 (.000***)<br>254 | -.186 (.000***)<br>254 | .427 (.000***)<br>251 | .436 (.000***)<br>249 | ----- |
| HEAD UP<br>N = | -.097 (.123)<br>254 | .062 (.326)<br>254 | -.214 (.001**)<br>251 | -.235 (.000***)<br>249 | ----- |
| TALK<br>N = | -.117 (.064)<br>251 | .304 (.000***)<br>251 | -.644 (.000***)<br>251 | -.628 (.000***)<br>249 | ----- |
| SOUND<br>N = | -.075 (.248)<br>242 | .370 (.000***)<br>242 | -.388 (.000***)<br>241 | -.379 (.000***)<br>239 | ----- |
| SMILE<br>N = | -.086 (.185)<br>240 | .313 (.000***)<br>240 | -.430 (.000***)<br>237 | -.420 (.000***)<br>235 | .485 (.000***)<br>237 |
| NEUTRAL<br>N = | .142 (028*)<br>240 | -.238 (.000***)<br>240 | .395 (.000***)<br>237 | .409 (.000***)<br>235 | -.285 (.000***)<br>237 |
| SOUND<br>N = | -.075 (.248)<br>242 | .370 (.000***)<br>242 | -.388 (.000***)<br>241 | -.379 (.000***)<br>239 | .533 (.000***)<br>241 |

*** Correlation is significant at the 0.001 level (2-tailed); ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed).

Head movement was correlated with negative valence, high arousal, off-task behavior and non-desirable states. This implies that students move their heads when they feel negative emotions, when being off-task and in a non-desirable learning state. When students are in such unproductive learning state, and when they are off-task, they tend to move their heads to the side. Also, students tend to move their head to the side when they have negative feelings. It is possible that students avoid the computer screen when they don't feel good about the software or the learning situation. At the same time, having their head in the middle had the opposite effect: it was correlated with positive valence, low arousal, on-task behavior, and desirable state for learning.

Students holding their head up indicates off-task behavior and an undesirable state for learning, while holding their head down is not (possibly because many students tend to work on paper on their desk). Again, head up could be an indication of screen avoidance. Talking and environmental sound are both correlated to high arousal and positive emotion, although they are associated with off-task behavior and undesirable states. This means that students tend to have off-task talk, which seems reasonable for a system that does not encourage on-task collaboration with a partner.

It seems obvious that frowning is related to having a negative valence emotion. However, frowning doesn't appear to be a good predictor of being on-task or being in a desirable learning state (not shown). A smile on the face does predict off-task behavior (R=-.430 with on-task) and undesirable state for learning (R=-.420), Table 3. Surprisingly, smiling was not linked to valence, but it is positively correlated with arousal and talk (students probably moved and talked with friends while they smiled). The opposite effect happened for a neutral face: it was positively correlated to desirable learning state and on-task behavior. A neutral face was linked to positive valence, most likely because we coded seeing a neutral emotion as positive valence. A neutral face was an indicator that the student was not moving (negative arousal) and not talking. Last, an environmental sound that is louder than background noise was a good predictor of talking (R=.533) suggesting that a microphone that senses for odd sounds can detect if a student is talking with good accuracy, which in turn was evidence for a non-desirable state for learning within the software, see Table 3.

## 5 Sensor Technology

**Figure 1**. Pressure mouse sensor

These human observations in the classroom are continuing as a way to understand the impact of student emotions on learning. Yet these student emotions can be detected automatically by intelligent tutors, which can then also respond dynamically with appropriate interventions. In order to establish a social and emotional connection with students, tutors should recognize students' affect and respond to them at a deep level. Towards this end, our goal in the second phase was to automate the observation process using sensors. We have developed a low cost multi-modal sensor platform that is being integrated into the Wayang Tutor and evaluated in classrooms. The platform includes a custom produced Pressure Mouse, a Wireless BlueTooth Skin Conductance sensor, a Posture Analysis Seat, and a Facial Expression System. This platform expands on an earlier one at an order of magnitude reduction in the overall cost. The sensors are developed from an earlier system that had several sensors in common with AutoTutor [9]. Pre-production prototypes of each sensor have been developed and we are producing thirty sets of these sensor platforms for simultaneous use in classrooms. The intent is to provide a better understanding of student behavior and affect and to determine the contribution of each sensor to the modeling of affect [14].

*Pressure mouse*. A pressure mouse is used to detect the increasing amounts of pressure users place on their mice related to their increased levels of frustration. The pressure mouse system has six force sensitive resistor sensors and an embedded microprocessor, Figure 1. It uses the standard communication channel of a USB mouse for pointing and clicking functions and then in parallel uses a second channel, a serial communications port, to provide pressure data at 20ms intervals from each of the six sensors. Pressure sensors located under the mouse button measure the force of the users click in addition to their overall pressure across the surface of the mouse.

*Posture Analysis Seat*. We have developed and are now testing a low-cost/low resolution pressure sensitive seat cushion and back pad with an incorporated accelerometer to measure elements of a student's posture and activity, Figure 2. This system captures many student movements relevant to education that were previously captured by the TekScan system, that used an extremely expensive Posture Analysis Seat, developed for medical and automotive applications [19]. The previous system used pattern recognition techniques while watching natural behaviors to *learn* which behaviors tended to accompany states such as interest and boredom. We are now developing similar algorithms based on the new low-cost posture analysis chair.

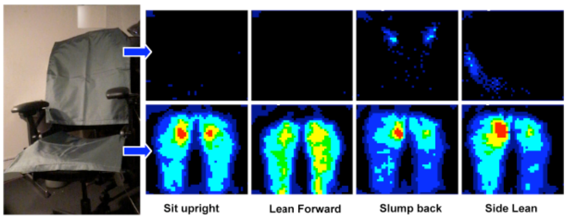


**Figure 2**. Posture State Chair Sensor. The previous sensor resulted in posture recognition (89-97% accurate). And classification of high/low interest and break taking (69-83% accurate) [14].

*Wireless skin conductance*. A wireless version of a earlier glove that sensed conductance was developed by Carson Reynolds and Marc Strauss at the MIT Media Lab, in collaboration with Gary McDarby, at Media Lab Europe, see Figure 3. While the skin conductance signal is not valenced (i.e. does not describe how positive or negative the affective state is) it is strongly correlated with arousal. High levels of arousal tend to accompany significant and attention-getting events [6].

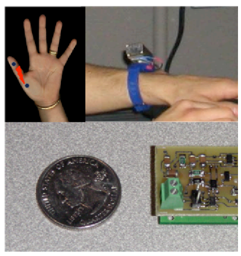

**Figure 3.** Wireless Skin Conductance Sensor

*Facial Expression Camera*. A person's mental state is not directly available to an observer; instead it is inferred from a range of non-verbal cues including facial expressions. We are using a facial expression recognition system that incorporates a computational model of mind reading as a framework for machine perception and mental state recognition [12]. This facial action analysis is based on a combination of bottom-up vision-based processing of the face (e.g. head nod or smile) with top-down predictions of mental state models (e.g. interest and confusion) to interpret the meaning underlying head and facial signals over time [12]. A multilevel, probabilistic architecture (using dynamic Bayesian networks) mimics the hierarchical manner with which people perceive facial and other human behavior [21] and handles the uncertainty inherent in the process of attributing mental states to others. The output probabilities represent a rich modality that technology can use to represent a person's state and respond accordingly. The resulting visual system infers mental states of people from head gestures and facial expressions in a video stream in real-time. At 30 fps, the inference system locates and tracks 24 feature points on the face and uses motion, shape and color deformations of these features to identify 20 facial and head movements (e.g., head pitch, lip corner pull) and 11 communicative gestures (e.g., head nod, smile, eyebrow flash) [21]. Dynamic Bayesian networks model these head and facial movements over time, and infer the student's "hidden" affective-cognitive state.

## 6 Discussion and Future Work

This paper described the use of human observations and wireless sensors to detect student emotions, learning, and attitudes towards learning. We identified emotion indicators (valence and arousal) that combined with on and off-task variables to represent desirable/undesirable states linked with student learning, as well as physical behaviors linked to emotional states. This was achieved through quantitative field observations in the classroom in which researchers recorded the behaviour of students using intelligent tutors. We described correlations between low-level observations (i.e. chair movement) and higher-level observations (valence, arousal, on-off task behavior) and then between these higher-level observations and student learning and attitudes. Through these links, we propose that low-level sensor information can tell us about emotion indicators and other state-variables linked to learning. Sensors can provide information about how students perform and information about when students are in non-productive states so that the tutor can provide appropriate interventions. In turn, sensors can also inform us whether the given interventions are working or not. With this goal in mind, low cost portable sensors are being used in natural classroom settings. Thus, once we know which variables are useful predictors of learning and affective outcomes, these sensors can replace the human observers and predict students' emotional states related to learning.

This paper unveiled several interesting findings: 1) observed fluctuating states of emotion and on/off task behavior help predict posttest performance and attitudes/motivation; 2) student states are expressed with specific behaviors that can be automatically detected with sensors; and 3) a mechanism for strong/weak learning behavior detection was identified. As a result of these findings we identify how sensors can predict and reflect student learning, see Table 4. Moving from right to left sensor readings and emotion/biologic indicators are used to predict student learning and other motivational variables; moving from left to right indicates how strong/weak learning and attitudes are expressed and detected by sensors.

**Table 4.** Guide to interpreting sensor data and predicting learning

| ⇐⇐ **Predict student learning and attitude** ⇐⇐ | | | |
|---|---|---|---|
| **Desirable Learning States** | **Emotion/task indicators** | **Biologic indicators** | **Sensors to use** |
| **Most desirable** (Joy, Aha moment, Concentrated Actively engaged) | + Valence AND On-task | Lean on hand; Little chair/head movement; Sit in middle of chair; Head in middle; Neutral face; | Chair sensors Camera |
| **Medium desirable** (Frustrated, angry) | -- Valence + Arousal | Head movement; Chair movement; Squeezing of mouse | Camera, Pressure mouse; Chair Sensors |
| **Least desirable** (Bored, tired) | Off-task OR -- Valence -- Arousal | Talking; Large chair movement; Head movement; Head to side or head up; Smile | Skin conductance; Camera; Chair sensors; Microphone |
| ⇒⇒ **Detect strong and weak student learning behavior** ⇒⇒ | | | |

Future work consists of using these behaviors to predict emotions and desirable/undesirable learning states that would in turn help us predict learning and attitudes towards learning mathematics. The long-term goal is to dynamically collect information about students' emotional state and predict student states, and in turn predict posttest performance in real time. Moreover, because certain states such as negative valence and high levels of arousal are unproductive for post-tutor assessments of learning/attitudes, such states will lead to the selection of an intervention. At that point we must also decide which interventions are most successful for individual students and context (e.g. topic, emotional state). Finally, we intend to resolve the nature of data from different sensors. The camera provides very high-level judgments as it uses its own inference engine to decide emotional states, whereas all other sensors provide relatively raw data. We are engaged in the development of machine learning algorithms that relate these data sets to learners' diverse emotional states. Using all of these techniques, we plan to recognize and help students cope with states of negative valence and support their return to on-task behavior.

**Acknowledgement.** This material is based upon work supported by the National Science Foundation, IIS Grant #0705554. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

1. Aleven, Vincent; Roll, Ido; McLaren, Bruce; Ryu, Eun Jeong; Koedinger, Ken. (2005). An architecture to combine meta-cognitive and cognitive tutoring: Pilot testing the Help

Tutor. *Proceedings of the 12th International Conference on Artificial Intelligence in Education*. 2005

2. Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Meheranian, H., Fisher, D., Barto, A., Mahadevan, S., Woolf, B.P. (2007) Repairing Disengagement with Non-Invasive Interventions. To appear in Proceedings of the 13th International Conference of Artificial Intelligence in Education. IOS Press.
3. Baker, R.S.J.d. (2007) Modeling and Understanding Students' Off-Task Behavior in Intelligent Tutoring Systems. Proceedings of ACM CHI 2007: Computer-Human Interaction.
4. Baker, R.S.J.d., Corbett, A.; and Koedinger, K. (2004). Detecting Student Misuse of Intelligent Tutoring Systems. *In Proceedings of the Seventh International Conference on Intelligent Tutoring Systems* 531–540.
5. Block, J. (1995). On the relation between IQ, impulsivity and deliquency. *Journal of American Psychology, 104*, 395-398.
6. Boucsein, W. (1992). Electrodermal activity. New York: Plenum Press.
7. Conati C. (2002) ‘Probabilistic assessment of users’ emotions in educational games,’ *Journal of Applied Artificial Intelligence, special issue on Merging Cognition and Affect in HCI”* 16(7-8):555–575.
8. Cytowic, R. E. (1989). *Synesthesia: A union of the senses*. New York: Springer Verlag.
9. D’Mello, S.K., Picard, R., & Graesser, A.C. (2007). Toward an affect-sensitive AutoTutor. IEEE Intelligent Systems, 22(4), 53-61.
10. Dweck, C. S. (1999). Self-theories: Their role in motivation, personality and development. Philadelphia : The Psychology Press
11. Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. New York: Bantam.
12. Kaliouby, R. and Robinson, P. (2005) “Real-time Inference of Complex Mental States from Facial Expressions and Head Gestures”, Real-Time Vision for HCI, chapter pages 181-200 Spring-Verlag, 2005.
13. Kapoor A, Mota S and Picard R W. (2001). Towards a learning companion that recognises affect, AAAI Fall Symposium 2001, North Falmouth, MA.
14. Kapoor, A., Picard, R.W., and Y. Ivanov (2004), "Probabilistic Combination of Multiple Modalities to Detect Interest," International Conference on Pattern Recognition, August 2004, Cambridge, U.K.
15. McQuiggan, S. and Lester, J., (2006) Diagnosing Self-Efficacy in Intelligent Tutoring Systems: An Empirical Study, In *Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (ITS 2006), Jhongli, Taiwan, June 26-30, 2006.
16. Norman, D. A. (1981). Twelve issues for cognitive science. In *Perspectives on cognitive science* (pp. 265-295). Hillsdale, NJ: Erlbaum.
17. Sheldon-Biddle E, Malone L and McBride D. (2003) ‘Objective measurement of student affect to optimize automated instruction’, Proceedings of Workshop on Modelling User Attitudes and Affect, *User Modeling ’03*.
18. Shute, V. J. (2006). Focus on formative feedback. ETS Research Report, #RR-07-11 Princeton, NJ
19. Tekscan (1997). Tekscan Body Pressure Measurement System User’s Manual. Tekscan Inc., South Boston, MA, USA.
20. Wundt, W. (1902) Outlines of Psychology. 2nd rev. English ed. London: Williams &Norgate.
21. Zacks., J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. Journal of Experimental Psychology: General, 130, 29-58.